

FIG. 1

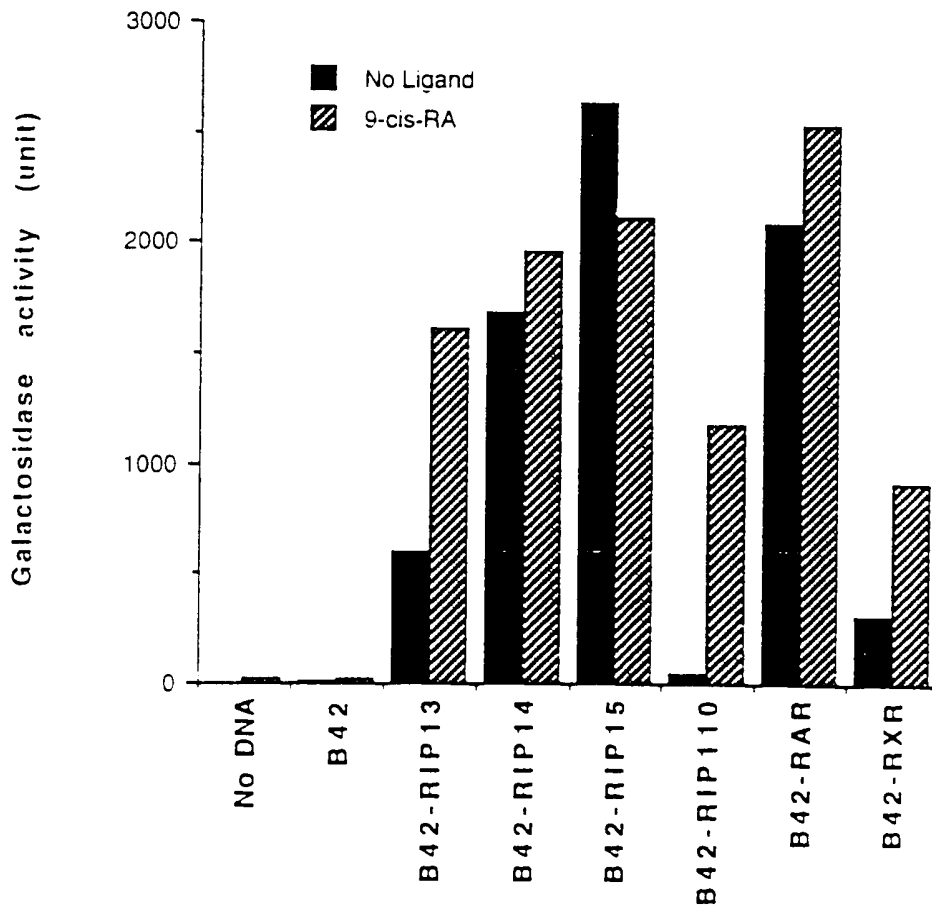


FIG. 2

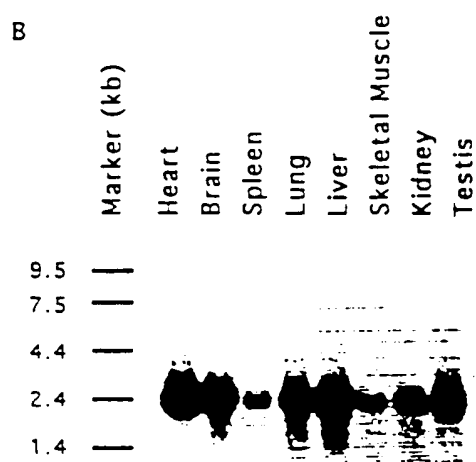
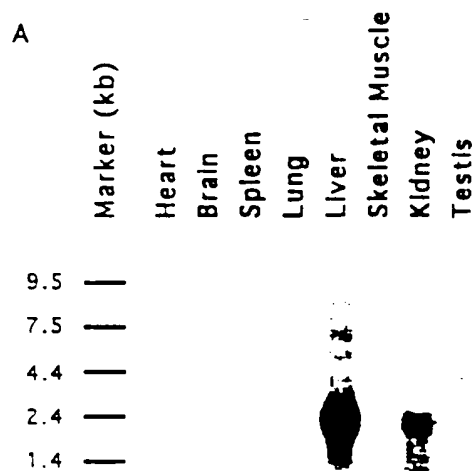


FIG. 3

Yeast clone

No. 6

No. 15

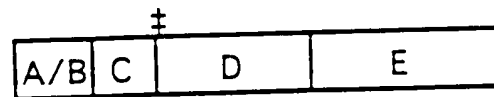
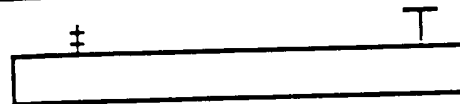
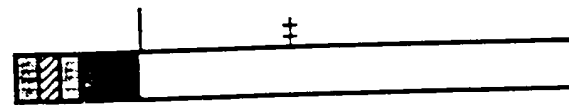
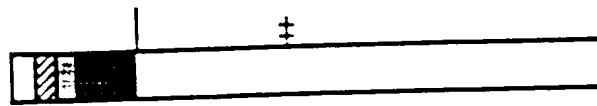
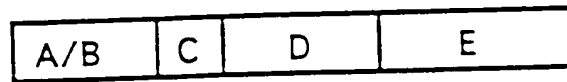
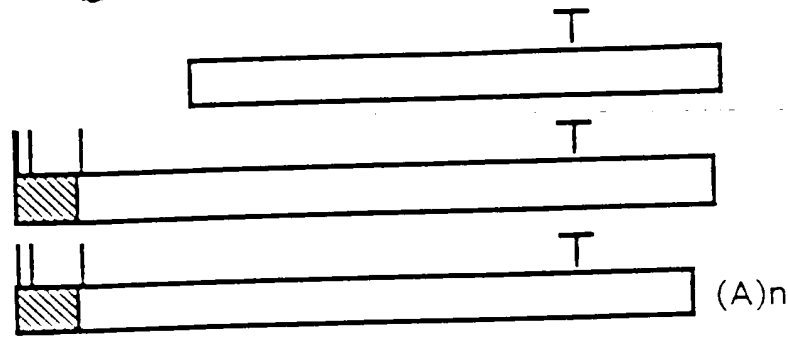
RIP14-1

No. 3

No. 12

No. 19

RIP14-2



- | ATG codon
- T Termination codon
- ‡ 12 bp addition
- ▨ 63 bp repeat

RIP14 Clone No

Specific to RIP14-2 clone 3

TGGTCAACCCAGGCTTGCGGTTCAGTATCTCCTCTCTTTCTCTCTCTAAGCCCATGTCTCTCT 60
CATATCACCCC 71

Specific to RIP14-2 clone 12

[illegible]

Common to RIF14-2

AGAGAAGAACCGAGTTTGTGAGAGTCTACAGCAAAATTTATCTGGGGCTATAGAAAGGAAAGT	131	128	-
GGGCTCCGAATCCCTCTATAGAGCGGTTGGACATCTCTGGGCCAAAGCAATCCAGGATCTTA	191	188	-
TTTGAGGACCAACCATCCCAAGTAGTACTTTCTCAGAGTTTGA-AAAGTTGGAGTGGTAGCCAA	251	248	-
GATGAATCTGATTGGGGCACTCCATTATACAGGCTACGGACGAGTTTTCTCTTTCTGAAAGC	311	308	-
TTAATT	317	314	-

Specific to RIP14-1

GGAGCTAAGGATGGTGATGCAGTTTCAGSSCTTAGAAAATCGAATTCAGATTAGTCTTC	-	-	60
M V M Q F Q G L E N P I Q I S L	-	-	16
ACCACAGCCACCGGCTGTCAGGATTGTTGTGCCGGACGGGATGAGTGTGAAGCCAGCTAAA	-	-	119
H H S H R L S G F V P D G M S V K P A K	-	-	36

Common to both isoforms

GGTATGCTAACAGAACACGGCGGCAGGCCCTCTGCGGGCAGATCTCGGATTTCGGAATCGTACT	378	375	180
G M L T E H A A G P L G Q N L D L E S Y	19	19	56
CCCCATACAACAATGTCCCGTTTCTCTCAAGTTCAGCCACAGATTTCCTCTCTCGTCTTACT	438	435	240
S P Y N N V P F P Q V Q P Q I S S S S Y	39	39	76
ATTCCGACCTGGGCTTCTACCCCCAACACCGGAGAACTGGTATTCTCTCTGGCATCTATG	498	495	300
Y S N L G F Y P Q Q P E D W Y S P G I Y	59	59	96
AACTCAGGGCATGCCCCGCTGAGACTGGGTACCCAGGGAGAACTGAGGTATCAGAGATGC	558	555	360
E L R R M P A E T G Y Q G E T E V S E M	79	79	116
CTGTGACAAAGAGCCCGCATGSCCGCGGCACTGGCAGGCAGATTAAGGGGGATGAGC	618	615	420
P V T K K P R M A A A S A G R I K G D E	99	99	136
TGTGTGTTGTCTGTGGAGACAGGGGCTCTGTGGGTACCACTACACCGGCTCACCTGTGAGG	678	675	480
L G N N C G D E A S G Y F Y N A L T C E	119	119	156

GCTGCAAGGTTTCTTTCGGAGAGCATTACCAAGACGGCGTGTACAGTGTAAAGAACG	738	735	540
G C K G E F R P S I T V N A V Y K C K N	139	139	176
GGGGCAACTGGGTGATGGACATGTACATGGCGCAGGAAGTGCCAGGAGTGCCGGCGTAAGGA	798	795	600
G G N C V M D M Y M R R K C Q E C R L R	159	159	196

GTATGTATACAG
M Y T G

RIP14-2 specific

AGTGCAGAGAGATGGGGATGTTGGGCTGAAT GTTGTGTAAGTGAATCCAGTGTAAATCTA	870	867	660
F C P E M G M L A E C L L T E I Q C K S	183	183	216
AACGGCTAAGGAAAATGTGAAGCAGCAGCGCTGNTCAGACAGTGAATGAGGACGACAGCG	930	927	720
K R L R K N V K Q H A D Q T V N E D D S	203	203	236
AAGGCGTGAATTGCGACAGTGAACCTCCACACCAAGTTTTGCAGGGAGAAAACGGAAC	990	987	780
E G R D L R Q V T S T T K F C R E K T E	223	223	256
TCACGGCAGACCAGCAGACCCCTCCTGGATTATATTTATGGATTGCTACAACAAACAGAGAA	1050	1047	840
L T A D Q Q T L L D Y I M D S Y N K Q R	243	243	276
TGCGTCAGGAAATCACAATAAAATCTTTAAAGAGAAATTTAGTGCAGAGAAAATTTTTC	1110	1107	900
M P Q E I T N K I L K E E F S A E E N F	263	263	296
TCATATTACAGAAATGGCAACCAGCCATGTACAGATTCTCGTAGAATTCACAAAAAGC	1170	1167	960
L I L T E M A T S H V Q I L V E F T K K	283	283	316
TTCCAGGGTTTCAGACACTGGACCCAGAGATCAGATTGCTTTTGTCAAAGGGTCCGCAG	1230	1227	1020
L P G F Q T L D H E D Q I A L L K G S A	303	303	336
TGGAGGCCATGTTTCTTCTGTTCCGGGGAGATTTTTCATAGAACTTCTGCGCGACATG	1290	1287	1080
V E A M F L R S A E I F N K K L P A G H	323	323	356
CAGACCTGTTGGAAGAAAGAAATTCGAAAGAGTGGTATCTCTGTGATGATATATAACCCCGA	1350	1347	1140
A D L L E E R I R K S G I S D E Y I T P	343	343	376
TGTTTCAGTTTCTATAAAAGTGTGGAGAACTCAAATGACTCAGGAGGAGTACGCTCTGC	1410	1407	1200
M F S F Y K S V G E L K M T Q E E Y A L	363	363	396
TCACAGCGATCGTCATCCTCTCTCCAGACAGACAATACATCAAGGACAGAGAGGGCGGTGG	1470(1453)	1260	
L T A I V I L S P D R Q Y I K D R E A V	383 (378)	416	
AGAAGCTGCAGGAGCCCCCTGCTTGATGTGCTACAAAAGCTGTGCAAGATGTACAGCCCTG	1530	-	1320
E K L Q E P L L D V L Q K L C K M Y Q P	403	-	436
AGAACCACAGCATTTTGGCGCTGCTCCTGGGTGCGCTGACGGAAGTCCGGACATTCAACC	(1554)	-	1380
E N P Q H F A C L L G R L T E L R T F N	(411)	-	456
ATCACCACCGTGAGATGCTGATGTCTTGGAGAGTGAATGATCAGAGTTTCACCCCGGTCC	-	-	1440
H H H A E M L M S W R V N D H K F T P L	-	-	476
TCTGTGAGATCTGGGATGTGCASTGATGGACACCACTGGGGGTGGTCCCTGTCTCTCTC	-	-	1500
L C E I W D V Q *	-	-	484

GGAACAGAAACCTTGTTCGTTTGTACCTGGTTTCACTCAGATCTCAATGAATATTTA	-	-	1560
TGTGGCAATTATACACCTCCACGGTTGTAAATACAGACTAGATAGAACTGCTTTCCCCA	-	-	1620
CACTGTATTTTACAAGGCTTCAGGAAACCCCACTGGCATGCCCTTTTGGCCTAATTAAAT	-	-	1680
CAATTGTTACTTCAATTCTATCTACTGAGCTAGGGGCATATTATTCTTCATTTCGACAATA	-	-	1740
TTATATATATTTTATAAAGTTGAGCTGTTTTCAACTGAGACAATAAA	-	-	1787

GCCAGGGCAACAGAGTCGGAGACCCCCCTGCCACCCCCCTCCCGATCGCOGGTGCAGTCAT -71
 GAGCCCCGGCTCCCCCTGGTGCACGGAGAGGGGGGGGGCTGGAACAAGCAGGCTGCTTC -11
GTCACCCACTATGTCTTCCCCCACAAGTTCTCTGGACACTCCCGTGCCTGGGAATGGTTC 50
 M S S P T S S L D T P V P G N G S 17
 TCCTCAGCCCCAGTACCTCCGCCACGTCACCCACTATTAAAGGAAGAGGGGCAGGAGACTGA 110
 P Q P S T S A T S P T I K E E G Q E T D 37
 TCCTCCTCCAGGCTCTGAAGGGTCCAGCTCTGCGTACATCGTGGTCACTTTAGAGCCAGA 170
 P P P G S E G S S S A Y I V V I L E P E 57
 GGATGAGCCTGAGCGCAAGCGGAAGAAGGGGCGGGCCCGAAGATGCTGGGCCATGAGCT 230
 D E P E R K R K K G P A P K M L G H E L 77
 GTGCCGGCTGTGCGGAGACAAGGCTTCGGGGCTTCCACTACAACGTGCTCAGCTGTGAAGG 290
C R V C G D K A S G F H Y N V L S C E G 97
 CTGCAAAGGCTTCTTCCGGCGCAGTGTGGTCCACGGTGGGGCGGGCGCTATGCCTGTCC 350
C K G F F R R S V V H G G A G R Y A C R 117
 GGGCAGCGGAACCTGCCAGATGGATGCCCTTCATGCGGCGCAAGTGCCAGCTCTGCCGGCT 410
G S G T C Q M D A F M R R K C C L C R L 137
 GCGCAAGTGCAAGGAGGCTGGCATGCGGGAGCAGTGGTGGTCTCTGAGGAGCAGATTCCG 470
R K C K E A G M R E Q C V L S E E Q I R 157
 GAAGAAAAGGATTTCAGAAGCAGCAACAGCAGCAGCCACCCCACTCTGAGCCAGCAGC 530
 K K R I Q K Q Q Q Q Q P P P P S E P A A 177
 CAGCAGCTCAGGCGCGGCCAGCGGCTCCCTGSCACTTCGGAAGCAAGCAGCCAGGGCTC 590
 S S S G R P A A S P G T S E A S S Q G S 197
 CGGGGAAGGAGAGGGCATCCAGCTGACCGCGGCTCAGGAGCTGATGATCCAGCAGTTAGT 650
 G E G E G I Q L T A A Q E L M I Q Q L V 217
 TGCCGGCGCAGCTGCAGTGCAACAAACGATCTTTCTCCGACCAGCCCCAAAGTCACGCGCTG 710
 A A Q L Q C N K R S F S D Q P K V T P W 237
 GCGGCTGGGTGCAGACCCCTCAGTCCCGAGATGCGCGTCAAGCAAGGCTTTGCCCCACTTCAC 770
 P L G A D P Q S R D A R Q Q R F A H F T 257
 CGAGCTAGCCATCATCTCGGTCCAGGAGATTGTGGACTTTGCCAAGCAGGTGCCAGGGTT 830
 E L A I I S V Q E I V D F A K Q V P G F 277
 CTTGCASTTGGGCGGGGAGGACCAGATCGCCCTCCTGAAGGCGTCCACCATTGAGATCAT 890
 L Q L G R E D Q I A L L K A S T I E I M 297
 GTTGTAGAAACAGCCAGAGCGTACAACCCAGAGACAGATGCATCAGTTCTGTAAGGA 950
 L L E T A R R Y N H E T E C I T F L K D 317

FIG. 6

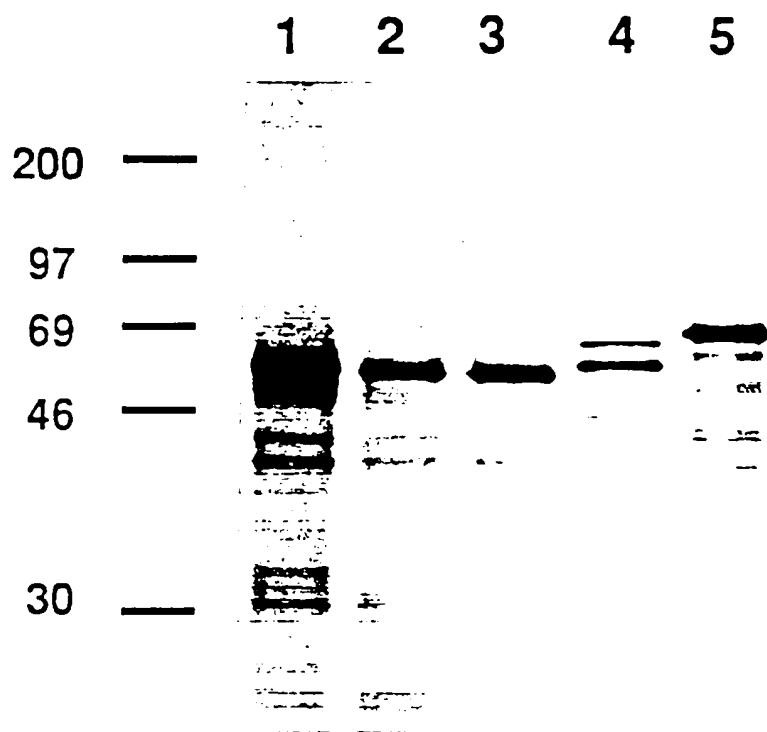


FIG. 7

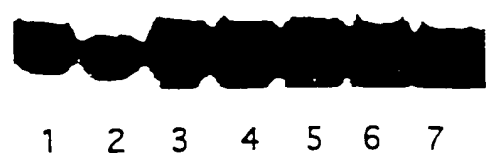
A.

RIP14		DNA		LIGAND/DIMER
EcR		82		37
RIP15		67		42
TR		61		26
RAR		59		31

B.

RIP15		DNA		LIGAND/DIMER
EcR		64		42
RAR		56		38
RXR		51		27
TR		50		32

		RIP14-1				cell lysate	
RXR	-	-	+	+	+	-	+
Competitor	-	-	-	sp	ns	-	-



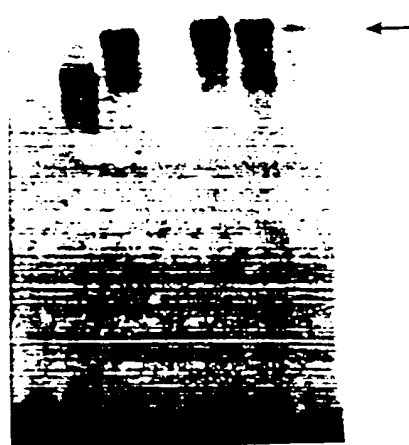
B

		RIP15				cell lysate	
RXR	-	-	+	+	+	-	+
Competitor	-	-	-	sp	ns	-	-



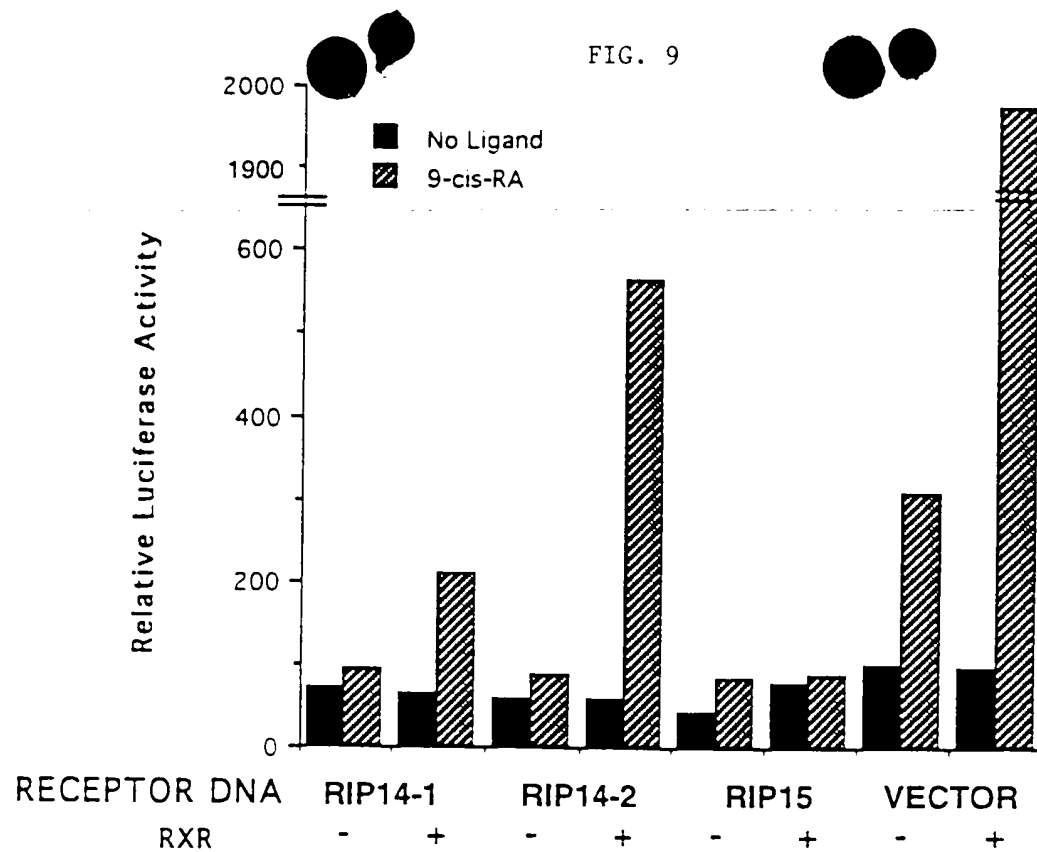
1 2 3 4 5 6 7

RIP14-1	-	+	+	+	+	-	-
RIP14-2	-	-	-	-	-	-	+
RIP14C	-	-	-	-	-	+	-
RXR	-	-	+	+	+	+	+
Competitor	-	-	-	sp	ns	-	-



8 9 10 11 12 13 14

FIG. 9



	TCTCCAAGTTGTGGCTTTCAGAGGAAGGATCTGGAACAACAGCAGTTGTGTCAAGTATAAAG	
12	-----+-----+-----+-----+-----+-----+-----+	71
	AGAGGTTCAACACCGAAAGTCTCCTTCCTAGACCTTTGTTTCGTCAACACAGTCATATTTTC	
c	S P S C G F Q R K D L E T S S C V S I K -	
	AAGAAGCGTAGACTTGAGGACTTACTCATAGTGATATCCGATAGCGATGGAGAGGAAACA	
72	-----+-----+-----+-----+-----+-----+-----+	131
	TTCTTCGCATCTGAACTCCTGAATGAGTATCACTATAGGCTATCGCTACCTCTCCTTTGT	
c	K K R R L E D L L I V I S D S D G E E T -	
	AAAGAGGAGAATGGATTGCAG/ AAACGAAGACAAAACAGTCGAACAGATCAAAGTGTTTG	
132	-----+-----+-----+-----+-----+-----+-----+	191
	TTTCTCCTCTTACCTAACGTCTTTTGCTTCTGTTTTGTCTAGTTTCACAAAC	
c	K E E N G L Q K T K T K Q S N R S K C L -	
	GCTAAAAGAAAAGTTGCACACATGTCAGAAGAAGAACAATTTGCTTTGGCTCTCAAGATG	
192	-----+-----+-----+-----+-----+-----+-----+	251
	CGATTTTCTTTTCAACGTGTGTACAGTCTTCTTCTTGTTAAACGAAACCGAGAGTTCTAC	
c	A K R K V A H M S E E E Q F A L A L K M -	
	AGTGAGCAGGAAGCTAGGGAGGTGAATAACCAGGAGGAGAAAGAAGAGGAGCTCTTGCGG	
252	-----+-----+-----+-----+-----+-----+-----+	311
	TCACTCGTCCCTTCGATCCCTCCACTTATTGGTCCTCCTCTTTCTTCTCCTCGAGAACGCC	
c	S E Q E A R E V N N Q E E K E E E L L R -	
	AAAGCCATTGCTGAAAGCCTGAATAGTTGCTGGTCTTCTGCTGCTTCTGCTACCAGATCT	
312	-----+-----+-----+-----+-----+-----+-----+	371
	TTTCGGTAACGACTTTTCGGACTTATCAACGACCAGAAGACGACGAAGACGATGGTCTAGA	
c	K A I A E S L N S C W S S A A S A T R S -	
	CGACCTCTGGCTGCTGAACTATCTTCACATTCCTCATCAAGAGAACACCAAAGACTCTGGG	
372	-----+-----+-----+-----+-----+-----+-----+	431
	GCTGGAGACCGACGACTTGATAGAAAGTGAAGGGTAGTTCTTCTGTGGTTTCTGAGACCC	
c	R P L A A E L S S H S H Q E N T K D S G -	
	ACCACTGAAGGCGTATGGCAGCTGGTACCTCCATCACTGTGTAAAGGCTCACATGTCAGT	
432	-----+-----+-----+-----+-----+-----+-----+	491
	TGGTGACTTCCGCATACCGTCGACCATGGAGGTAGTGACACATTTCCGAGTGTACAGTCA	
c	T T E G V W Q L V P P S L C K G S H V S -	
	CAGGGAAACGAGGCTGAGCAAAGAAAGGAGCCCTGGGAcCACAATGAAAACACTGAAGAG	
492	-----+-----+-----+-----+-----+-----+-----+	551
	GTCCCTTTGCTCCGACTCGTTTCTTTCCCTCGGGACCCTgGTGTACTTTTGTGACTTCTC	
c	Q G N E A E Q R K E P W D H N E N T E E -	
	GAGCCGGTCTCTGGCAGCTCAGGAAGCTGGGACCAGTCAAGCCAGCCAGTGTTTGAGAAT	
552	-----+-----+-----+-----+-----+-----+-----+	611
	CTCGGCCAGAGACCGTCGAGTCCCTTCGACCCTGGTCAAGTTCGGTCGGTCACAAACTCTTA	

C E P V S G S S G S W D Q S S Q P V F E N -
 GAGAACGTTAAATGTTTTGACAGATGTACTGGCCACTTGGCTGAGCACACACAGTGTGGG
 612 -----+-----+-----+-----+-----+-----+-----+ 671
 CTCTTGCAATTTACAAAAGTGTCTACATGACCGGTGAACCGACTCGTGTGTGTACACCC

C E N V K C F D R C T G H L A E H T Q C G -
 AAGCCACAGGAAAGTACTGGGAGTGGTTATGCTTTTTCCAAAGCTGTCCAGGGTAGGGGG
 672 -----+-----+-----+-----+-----+-----+-----+ 731
 TTCGGTGTCTTTTCATGACCCTCACCAATACGAAAAAGGTTTCGACAGGTCCCATCCCCC

C K P Q E S T G S G Y A F S K A V Q G R G -
 GACACGTCTAGGCAATGCCTTCCTATCCCAGCAGACACAAAAGGTCTCCAGGACACTGGG
 732 -----+-----+-----+-----+-----+-----+-----+ 791
 CTGTGCAGATCCGTTACGGAAGGATAGGGTCTGTCTGTGTTTTCCAGAGGTCTGTGACCC

C D T S R Q C L P I P A D T K G L Q D T G -
 GGCACGTGTGCACTACTACTGGGGTATTCCATTCTGCCCTGCTGGAGTAGATCCCAATCAA
 792 -----+-----+-----+-----+-----+-----+-----+ 851
 CCGTGACACGTGATGATGACCCCATAGGTAAGACGGGACGACCTCATCTAGGGTTAGTT

C G T V H Y Y W G I P F C P A G V D P N Q -
 TACACCAATGTCAATTCTCTGCCAGTTAGAGGTTTATCAGAAGAGCCTGAAAAATGGCTCAG
 852 -----+-----+-----+-----+-----+-----+-----+ 911
 ATGTGGTTACAGTAAGAGACGGTCAATCTCCAAATAGTCTTCTCGGACTTTTACCGAGTC

C Y T N V I L C Q L E V Y Q K S L K M A Q -
 AGACAGCTTGTTAAAAAAGAGGGTTTGGGGAACCAGTGTTACCTAGACCTCCTTTTCTG
 912 -----+-----+-----+-----+-----+-----+-----+ 971
 TCTGTGCAACAATTTTTTCTCCAAACCCCTTGGTCACAATGGATCTGGAGGAAAAGAC

C R Q L V K K R G F G E P V L P R P P F L -
 ATCCAGAATGAATGTGGCCAAGAAGATCAGACTAGTGACAAAAATGAAGGCATCTCAGAA
 972 -----+-----+-----+-----+-----+-----+-----+ 1031
 TAGGTCTTACTTACACCGGTTCTTCTAGTCTGATCACTGTTTTTACTTCCGTAGAGTCTT

C I Q N E C G Q E D Q T S D K N E G I S E -
 GATATGGGAGATGAAGCCAAAGAGGAAAGGCAGGAATCTAGGGCATCTGTCTGGCACTCA
 1032 -----+-----+-----+-----+-----+-----+-----+ 1091
 CTATACCTCTACTTCCGTTTCTCCTTTCCGTCCTTAGATCCCGTAGACAGACCGTGAGT

C D M G D E A K E E R Q E S R A S V W H S -
 GAAACCAAGGATTTTCAAAAAAGTCCAATTAAAGCTTGAAACAGAACTTTTGTGGAG
 1092 -----+-----+-----+-----+-----+-----+-----+ 1151
 CTTTGGTTCCTAAAAGTTTTTTCAGGTTAATTTTCGAACTTTGTCTTTGAAAACAACCTC

C E T K D F Q K S P I K S L K Q K L L L E -
 GAAGAACCAACAACCAGTCGTGGTCAGTCTTCCCAAGGTCTGTTTGTGTAAGAAACCTCT
 1152 -----+-----+-----+-----+-----+-----+-----+ 1211
 CTTCTTGGTTGTTGGTCAGCACCAGTCAGAAGGGTCCAGACAAACAACCTCTTTGGAGA

C E E P T T S R G Q S S Q G L F V E E T S -

C

[illegible]

GGATGCCCTGGCTGCTCTTGTGGATGCTGCAGCTTCTGCACCCCAGATGGATGTTTCCAA
 901 -----+-----+-----+-----+-----+-----+ 960
 CCTACGGGACCGACGAGAACACCTACGACGTCGAAGACGTGGGGTCTACCTACAAAGGTT
 b D A L A A L V D A A A S A P Q M D V S K -
 AACAAAAGAGAGTAAGCATGAAGCTGCCAGGTTAGAAGAAAATTTGAGAAGCAGGTCAGC
 961 -----+-----+-----+-----+-----+-----+ 1020
 TTGTTTTCTCTCATTCTGTAAGCGACGGTCCAATCTTCTTTTAACTCTTCGTCCAGTCG
 b T K E S K H E A A R L E E N L R S R S A -
 AGCAGTTAGTGAACAGCAGCAGCTAGAGCAGAAAAACCTGGAGGTGGAGAAGAGATCTGT
 1021 -----+-----+-----+-----+-----+-----+ 1080
 TCGTCAATCACTTGTCTGTCGTCGATCTCGTCTTTTGGACCTCCACCTCTTCTCTAGACA
 b A V S E Q Q Q L E Q K N L E V E K R S V -
 TCAGTGTGTGTGCACTTCTTCAGCCCTTCCAAGTGGCAAGGCCAGCCTCATGCCTCAGT
 1081 -----+-----+-----+-----+-----+-----+ 1140
 AGTCACACACACGTGAAGAAGTCGGGAAGGTTACCGGTTCCGGGTCCGAGTACGGAGTCA
 b Q C V C T S S A L P S G K A Q P H A S V -
 AGTGTATTCTGAGGCTGGGAAAGATAAAGGGCCTCCTCCAAAATCCAGATATGAGGAAGA
 1141 -----+-----+-----+-----+-----+-----+ 1200
 TCACATAAGACTCCGACCCTTTCTATTTCCCGGAGGAGGTTTTAGGTCTATACTCCTTCT
 b V Y S E A G K D K G P P P K S R Y E E E -
 GCTAAGGACCCGAGGGAAGACTACCATTACTGCAGCTAACTTCATAGACGTGACCATCAC
 1201 -----+-----+-----+-----+-----+-----+ 1260
 CGATTCCCTGGGCTCCCTTCTGATGGTAATGACGTCGATTGAAGTATCTGCACTGGTAGTG
 b L R T R G K T T I T A A N F I D V T I T -
 CCGGCAAATTGCCTCGGACAAGGATGCGAGGGAACGTGGCTCTCAAAGTTCAGACTCTTC
 1261 -----+-----+-----+-----+-----+-----+ 1320
 GGCCGTTTAAACGGAGCCTGTTCTACGCTCCCTTGCACCGAGAGTTTCAAGTCTGAGAAG
 b R Q I A S D K D A R E R G S Q S S D S S -
 TAGTAGCTTGTCTTCTCACAGGTATGAAACGGCTAGTGATGCCATTGAGGTGATAAGTCC
 1321 -----+-----+-----+-----+-----+-----+ 1380
 ATCATCGAACAGAAGAGTGTCCATACTTTGCCGATCACTACGGTAACTCCACTATTGAGG
 b S S L S S H R Y E T A S D A I E V I S P -
 CGCCAGCTCACCTGCACCACCCAGGAAAAGCCACAGGCCTATCAGCCAGACATGGTTAA
 1381 -----+-----+-----+-----+-----+-----+ 1440
 GCGGTTCGAGTGGACGTGGTGGGGTCTTTTTCGGTGTCCGGATAGTCGGTCTGTACCAATT
 b A S S P A P P Q E K P Q A Y Q P D M V K -

1441 GGCAAATCAAGCAGAAAAATGAGTCCACTCGACAGTATGAAGGTCCACTGCATCATTATCG 1500
 -----+-----+-----+-----+-----+-----+-----+
 CCGTTTAGTTCCTCTTTTACTCAGGTGAGCTGTCACTTCCAGGTGACGTAGTAATAGC
 b A N Q A E N E S T R Q Y E G P L H H Y R -
 1501 GTCCCAGCAGGAATCACCATCTCCACAGCAACAGCCACCACTGCCCCCATCTTCCCAGTC 1560
 -----+-----+-----+-----+-----+-----+-----+
 CAGGGTCCTCTTAGTGGTAGAGGTGTCGTTGTTCGGTGGTGACGGGGGTAGAAGGGTCAG
 b S Q Q E S P S P Q Q Q P P L P P S S Q S -
 1561 AGAGGGGAATGGGACAGGTGCCCAGGACCCATCGACTGATCACACTTGCTGACCACATCTG 1620
 -----+-----+-----+-----+-----+-----+-----+
 TCTCCCTTAACCTGTCCACGGTCCTGGGTAGCTGACTAGTGTGAACGACTGGTGTAGAC
 b E G M G Q V P R T H R L I T L A D H I C -
 1621 TCAAAATTATCACACAAGATTTTGCTAGAAATCAAGTTCCTCGCAGCCTTCTACTTCTAC 1680
 -----+-----+-----+-----+-----+-----+-----+
 AGTTTAATAGTGTGTTCTAAAACGATCTTTAGTTCAAGGGAGCCTCGGAAGATGAAGATG
 b Q I I T Q D F A R N Q V P S Q P S T S T -
 1681 ATTCCAAACTTCACCATCTGCTTTGTCATCCACACCTGTAAGAACTAAAACCTCAAGCCG 1740
 -----+-----+-----+-----+-----+-----+-----+
 TAAGSTTTGAAGTGGTAGACGAAACAGTAGGTGTGGACATTCTTGATTTTGGAGTTCCGGC
 b F Q T S P S A L S S T P V R T K T S S R -
 1741 CTACAGCCCAGAATCACAGTCTCAGACTGTCTTGCATCCCAGAACAGGTCTTAGAGTCTC 1800
 -----+-----+-----+-----+-----+-----+-----+
 GATGTCCGGTCTTAGTGTCCAGAGTCTGACAGAACGTAGGGTCTGGTCCAGGATCTCAGAG
 b Y S P E S Q S Q T V L H P R P G P R V S -
 1801 TCCAGAAAATCTTGTGGATAAATCCCGGGGAAGCAGGCCTGGAAAATCTCCAGAGAGGAG 1860
 -----+-----+-----+-----+-----+-----+-----+
 AGGTCTTTTAGAACACCTATTTAGGGCCCCCTTCGTCCGGACCTTTTAGAGGTCTCTCCTC
 b P E N L V D K S R G S R P G K S P E R S -
 1861 TCATATCCCATCAGAGCCCTATGAGCCCATCTCCCCACCCCAAGGCCCTGTGTGCATGA 1920
 -----+-----+-----+-----+-----+-----+-----+
 AGTATAGGGTAGTCTCGGGATACTCGGGTAGAGGGGTGGGGTTCCGGGACGACACGTACT
 b H I P S E P Y E P I S P P Q G P A V H E -
 1921 GAAGCAGGACAGCATGTTGCTCTTETCACAGAGGGGAGTGGACCTGTCTGAGCAAAGGAG 1980
 -----+-----+-----+-----+-----+-----+-----+
 CTTCTGCTCTGTCTACAACGAGAACAGTGTCTCCCTCACCTGGGACGACTCGTTTCCTC
 b K Q D S M L L L S Q R G V D P A E Q R S -
 1981 TGATTCTCGATCACCAGGAAGTATAAGCTACTTGCCTTCATTCTTCACCAAGCTTGAAAG 2040
 -----+-----+-----+-----+-----+-----+-----+
 ACTAAGAGCTAGTGGTCCTTCATATTTCGATGAACGGAAGTAAGAAGTGGTTTGAACCTTTC
 b D S R S P G S I S Y L P S F F T K L E S -
 CACATCACCCATGGTTAAATCAAAGAAACAGGAAATTTTTCGTAAGTTGAACTCTTCTGG

2041 -----+-----+-----+-----+-----+-----+-----+ 2100
 GTGTAGTGGGTACCAATTTAGTTTCTTTGTCTTTAAAAAGCATTCAACTTGAGAAGACC
 b T S P M V K S K K Q E I F R K L N S S G -
 TGGAGGTGACTCTGATATGGCAGCTGCTCAGCCAGGAACAGAGATCTTCAATCTGCCAGC
 2101 -----+-----+-----+-----+-----+-----+ 2160
 ACCTCCACTGAGACTATACCGTCGACGAGTCGGTCCTTGTCTCTAGAAGTTAGACGGTCG
 b G G D S D M A A A Q P G T E I F N L P A -
 AGTTACCACATCAGGTGCAGTGAGCTCAAGAAGCCATTCTTTTGTCTGATCCCGCCAGTAA
 2161 -----+-----+-----+-----+-----+-----+ 2220
 TCAATGGTGTAGTCCACGTCAGTCTCGAGTCTTCCGTAAGAAAACGACTAGGGCGGTCAATT
 b V T T S G A V S S R S H S F A D P A S N -
 CCTTGGTCTAGAAGACATCATCAGAAAGGCTCTCATGGGAAGTTTTGATGATAAAGTTGA
 2221 -----+-----+-----+-----+-----+-----+ 2280
 GGAACCAGATCTTCTGTAGTAGTCTTCCGAGAGTACCCTTCAAACTACTATTTCAACT
 b L G L E D I I R K A L M G S F D D K V E -
 AGATCATGGTGTGTGTATGTCCCATCTGTGGGCATTATGCCTGGTAGTGCCAGCACCTC
 2281 -----+-----+-----+-----+-----+-----+ 2340
 TCTAGTACCACAACAGTACAGGGTAGGACACCCGTAATACGGACCATCACGGTCGTGGAG
 b D H G V V M S H P V G I M F G S A S T S -
 AGTGGTGACGAGCAGCGAGGCACGGAGAGATGAAGGGGAGCCATCACCTCATGCAGGAGT
 2341 -----+-----+-----+-----+-----+-----+ 2400
 TCACCACTGCTCGTCGCTCCGTGCCTCTCTACTTCCCCTCGGTAGTGGAGTACGTCTCTCA
 b V V T S S E A R R D E G E P S P H A G V -
 ATGCAAACCAAAGCTGATCAACAAATCAAACAGCAGGAAGTCTAAATCTCTATTCCTGG
 2401 -----+-----+-----+-----+-----+-----+ 2460
 TACGTTTGGTTTCGACTAGTTGTTTAGTTTGTCTCTTCAGATTTAGAGGATAAGGACC
 b C K P K L I N K S N S R K S K S P I P G -
 GCAAAGCTATTTAGGAACTGAAAGGCCTTCTTCTGTCTCCTCTGTGCATTGAGAAGGTGA
 2461 -----+-----+-----+-----+-----+-----+ 2520
 CGTTTCGATAAATCCTTGACTTTCCGGAAGAAGACAGAGGAGACACGTAAGTCTTCCACT
 b Q S Y L G T E R P S S V S S V H S E G D -
 TTACCACAGGCAGACACCAGGATGGGCATGGGAAGATCGGCCCTCTTCAACAGGTTCTAC
 2521 -----+-----+-----+-----+-----+-----+ 2580
 AATGGTGTCCGTCTGTGGTCCTACCCGTACCCTTCTAGCCGGGAGAAGTTGTCCAAGATG
 b Y H R Q T P G W A W E D R P S S T G S T -
 TCAGTTCCTTACAACCCTCTGACCATACGGATGCTCAGCAGTACACCACCTACACAGAT
 2581 -----+-----+-----+-----+-----+-----+ 2640
 AGTCAAGGGAATGTTGGGAGACTGGTATGCCTACGAGTCGTCATGTGGTGGATGTGTCTA
 b Q F P Y N P L T I R M L S S T P P T Q I -
 CGCATGCGCCCCATCTGCCATCACCCAAGCAGCTCCACATCAACAGAACCGCATCTGGGA
 2641 -----+-----+-----+-----+-----+-----+ 2700
 GCGTACGCGGGGTAGACGGTAGTGGGTTCGTGAGGTGTAGTTGTCTTGGCGTAGACCTT

b A C A P S A I T Q A A P H Q Q N R I W E -

2701 GAGGGAGCCTGCCCCGCTCCTCTCAGCGCAGTATGAGACACTGTCTGATAGTGACGACTG
 -----+-----+-----+-----+-----+-----+ 2760
 CTCCTCGGACGGGCGAGGAGAGTCGCGTCATACTCTGTGACAGACTATCACTGCTGAC

b R E P A P L L S A Q Y E T L S D S D D * -

2761 AGCTGTGCGTGGBAGAGCGCTCTGGCTTTGGTTTTTATTGAAGATTTAAAAAAAAAAAAA
 -----+-----+-----+-----+-----+-----+ 2820
 TCGACACGCACCTCTCGCGAGACCGAAACCAAAAATAACTTCTAAATTTTTTTTTTTTTT

AA
 2821 -- 2822
 TT